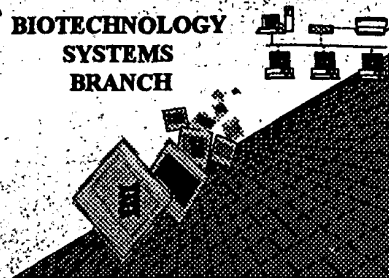


# **RAW SEQUENCE LISTING** **ERROR REPORT**

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number:

09/249,543

Art Unit / Team No:

01P6

Date Processed by STIC:

2/26/99

**THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.**

**PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:**

**1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,**

**2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY**

**THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.**

**IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:**

**ARTI SHAH 703-308-4212**



UNITED STATES DEPARTMENT OF COMMERCE  
Patent and Trademark Office

Address: COMMISSIONER OF PATENTS AND TRADEMARKS  
Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NO.
---------------	-------------	-----------------------	---------------------

EXAMINER
----------

ART UNIT	PAPER NUMBER
----------	--------------

3

DATE MAILED:

Commissioner of Patents

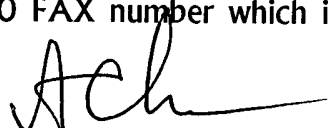
Please find below a communication from the EXAMINER in charge of this application.

This application contains sequence disclosures that are encompassed by the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR 1.821(a)(1) and (a)(2). A computer readable form (CRF) of the sequence listing was submitted. However, the CRF could not be processed by the Scientific and Technical Information Center (STIC) for the reason(s) set forth on the attached CRF Diskette Problem Report. Specifically, problems with designations required by the new rules are indicated in Item 10 of the attached Error Summary.

Applicant is given ONE MONTH, or THIRTY DAYS, whichever is longer, from the mailing date of this letter within which to comply with the sequence rules, 37 CFR 1.821 - 1.825. Failure to comply with these requirements will result in ABANDONMENT of the application under 37 CFR 1.821(g). Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136. In no case may an applicant extend the period for response beyond the SIX MONTH statutory period. Direct the reply to the undersigned. Applicant is requested to return a copy of the attached CRF Diskette Problem Report with the reply.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to William W. Moore whose telephone number is (703) 308-0583. The examiner can be reached Monday through Friday from 9:00 AM to 5:30PM EST. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Ponnathapura Achutamurthy, can be reached at (703) 308-3804. Papers related to this application may be submitted to Group 1800 by facsimile transmission. The faxing of such papers must conform with the notice published November 15, 1989 in the Official Gazette, 1096 OG 30. Informal and unofficial communications may be sent to the Art Unit 1652 FAX number, (703) 308-0294. Official filings should be sent to the Technical Center 1600 FAX number which is (703) 308-4556.

William W. Moore  
August 27, 1999

  
PONNATHAPUACHUTAMURTHY  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 1600

# **NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- ☐ 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- ☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- ☒ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- ☒ 7. Other: See attached Error Report and Error Summary

## **Applicant must provide:**

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- ☐ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123  
 For CRF submission help, call (703) 308-4212  
 For PatentIn software help, call (703) 308-6856

**Please return a copy of this notice with your response.**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/249,543

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2      Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3      Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.  
All text must be visible on page.
- 4      Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and uses spacing between the numbers.
- 5      Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6      Variable Length      Sequence(s)      contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) features section that some may be missing.
- 7      Wrong Designation      Sequence(s)      contain amino acid or nucleic acid designators which are not standard representations as per the Sequence Rules (Please refer to paragraph 1.822)
- 8      Skipped Sequences (OLD RULES)      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(1) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xl) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9      Skipped Sequences (NEW RULES)      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10      Use of N's or Xaa's (NEW RULES)      Use of N's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
- 11      Use of <213>Organism (NEW RULES)      Sequence(s)      are missing this mandatory field or its response.
- 12      Use of <220>Feature (NEW RULES)      Sequence(s)      are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
(Se "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)  
(Sec. 1.823 of new Sequence Rules)
- 13      Wrong Format      File submitted was in the alphabetical heading format of the Old Sequence Rules. This is invalid since the "Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures" Federal Register Notice, Vol. 63, No. 104, June 1, 1998, p. 29620 applies to applications filed on or after July 1, 1998.  
AKS-Biotechnology Systems Branch- 7/10/98

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/249,543DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

Does Not Comply

Corrected Diskette Needed

This Raw Listing contains the General Information  
Section and up to first 5 pages.

1 <110> APPLICANT: Evans, Thomas  
2 Xu, Ming-Qun  
3 <120> TITLE OF INVENTION: Intein-Mediated Protein Ligation Of Expressed Proteins  
4 <130> FILE REFERENCE: NEB-154  
5 <140> CURRENT APPLICATION NUMBER: US/09/249,543  
6 <141> CURRENT FILING DATE: 1999-02-12  
7 <160> NUMBER OF SEQ ID NOS: 24  
8 <170> SOFTWARE: PatentIn Ver. 2.0  
9 <210> SEQ ID NO 1  
10 <211> LENGTH: 99  
11 <212> TYPE: DNA  
12 <213> ORGANISM: Artificial Sequence  
13 <220> FEATURE:  
14 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
15 Synthesized From Methanobacterium  
16 thermoautotrophicum.  
17 <400> SEQUENCE: 1  
18 tcgaggcaac caaccctgc gtatccggtg acaccattgt aatgactagt ggcggtccgc 60  
19 gcactgtggc tgaactggag ggcaaaccgt tcaccgcac 99  
20 <210> SEQ ID NO 2  
21 <211> LENGTH: 93  
22 <212> TYPE: DNA  
23 <213> ORGANISM: Artificial Sequence  
24 <220> FEATURE:  
25 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
26 Synthesized From Methanobacterium  
27 thermoautotrophicum.  
28 <400> SEQUENCE: 2  
29 ccggttggt gctcgccaca gttgtgtaca atgaagccat tagcagtga tgcgctagca 60  
30 ccgtaaacag tagcgtcata aacatcctgg cgg 93  
31 <210> SEQ ID NO 3  
32 <211> LENGTH: 100  
33 <212> TYPE: DNA  
34 <213> ORGANISM: Artificial Sequence  
35 <220> FEATURE:  
36 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
37 Synthesized From Methanobacterium  
38 thermoautotrophicum.  
39 <400> SEQUENCE: 3  
40 tgattcgcg gctctggctac ccatgccct caggtttctt ccgcacctgt gaacgtgacg 60  
41 tatatgatct gcgtacacgt gagggtcatt gcttacgttt 100  
42 <210> SEQ ID NO 4  
43 <211> LENGTH: 100  
44 <212> TYPE: DNA

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/249,543DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

```
45 <213> ORGANISM: Artificial Sequence
46 <220> FEATURE:
47 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
48     Synthesized From Methanobacterium
49     thermoautotrophicum.
50 <400> SEQUENCE: 4
51     gacctatgat caccgtgttc tggatgatgga tggatggcctg gaatggcgtg ccgagggtga 60
52     actggaacgc ggcgaccgcc tggatgatgga tgatgcagct 100
53 <210> SEQ ID NO 5
54 <211> LENGTH: 87
55 <212> TYPE: DNA
56 <213> ORGANISM: Artificial Sequence
57 <220> FEATURE:
58 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
59     Synthesized From Methanobacterium
60     thermoautotrophicum.
61 <400> SEQUENCE: 5
62     ggcagatttc cggcactggc aaccttccgt ggcctgcgtg gcgctggccg ccaggatggt 60
63     tatgacgcta ctgtttacgg tgctagc 87
64 <210> SEQ ID NO 6
65 <211> LENGTH: 49
66 <212> TYPE: DNA
67 <213> ORGANISM: Artificial Sequence
68 <220> FEATURE:
69 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
70     Synthesized From Methanobacterium
71     thermoautotrophicum.
72 <400> SEQUENCE: 6
73     gcattcactg ctaatggctt cattgtacac aactgtggcg agcagccaa 49
74 <210> SEQ ID NO 7
75 <211> LENGTH: 100
76 <212> TYPE: DNA
77 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
80     Synthesized From Methanobacterium
81     thermoautotrophicum.
82 <400> SEQUENCE: 7
83     ccagcgccac gcaggccacg gaaggttgcc agtgccggaa actcgccagc tgcacatcc 60
84     atcaccaggc ggtcgccgcg ttccagttca cccgcggcac 100
85 <210> SEQ ID NO 8
86 <211> LENGTH: 90
87 <212> TYPE: DNA
88 <213> ORGANISM: Artificial Sequence
89 <220> FEATURE:
90 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
91     Synthesized From Methanobacterium
92     thermoautotrophicum.
93 <400> SEQUENCE: 8
94     gccattccag gccaccatcc atcaccagaa caccgtgatc atgggtcaaa cgtaagcaat 60
```

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/249,543

DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

```

95      gaccctcacg tgtacgcaga tcatatacgt                      90
96      <210> SEQ ID NO 9
97      <211> LENGTH: 97
98      <212> TYPE: DNA
99      <213> ORGANISM: Artificial Sequence
100     <220> FEATURE:
101     <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
102           Synthesized From Methanobacterium
103           thermoautotrophicum.
104     <400> SEQUENCE: 9
105           cacgttcaca ggtgcggaag aaacctgagg ggcattgggta gccagagccg cgaatcagtg 60
106           cggatgaacgg ttgacctcc agttcagcca cagtgcg                      97
107     <210> SEQ ID NO 10
108     <211> LENGTH: 55
109     <212> TYPE: DNA
110     <213> ORGANISM: Artificial Sequence
111     <220> FEATURE:
112     <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
113           Synthesized From Methanobacterium
114           thermoautotrophicum.
115     <400> SEQUENCE: 10
116           cggaccgcca ctatgcatta caatggtgtc accggatagc caggggttgg ttgcc      55
117     <210> SEQ ID NO 11
118     <211> LENGTH: 45
119     <212> TYPE: DNA
120     <213> ORGANISM: Artificial Sequence
121     <220> FEATURE:
122     <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
123           Synthesized From Methanobacterium
124           thermoautotrophicum.
125     <400> SEQUENCE: 11
126           tcgaggcaac caacgcattg gtatccggtg acaccattgt aatga              45
127     <210> SEQ ID NO 12
128     <211> LENGTH: 45
129     <212> TYPE: DNA
130     <213> ORGANISM: Artificial Sequence
131     <220> FEATURE:
132     <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
133           Synthesized From Methanobacterium
134           thermoautotrophicum.
135     <400> SEQUENCE: 12
136           ctatgcatta caatggtgtc accggatagc catgcgttgg ttgcc              45
137     <210> SEQ ID NO 13
138     <211> LENGTH: 36
139     <212> TYPE: DNA
140     <213> ORGANISM: Artificial Sequence
141     <220> FEATURE:
142     <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
143           Synthesized From Methanobacterium
144           thermoautotrophicum.

```

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/249,543DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

145 <400> SEQUENCE: 13  
146 tcgagggtcg cgtatccggt gacaccattg taatga 36  
147 <210> SEQ ID NO 14  
148 <211> LENGTH: 36  
149 <212> TYPE: DNA  
150 <213> ORGANISM: Artificial Sequence  
151 <220> FEATURE:  
152 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
153 Synthesized From Methanobacterium  
154 thermoautotrophicum.  
155 <400> SEQUENCE: 14  
156 ctagtcatca caatggtgac accggatacg cagccc 36  
157 <210> SEQ ID NO 15  
158 <211> LENGTH: 54  
159 <212> TYPE: DNA  
160 <213> ORGANISM: Artificial Sequence  
161 <220> FEATURE:  
162 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
163 Synthesized From Methanobacterium  
164 thermoautotrophicum.  
165 <400> SEQUENCE: 15  
166 tcgagggtcg cgaggcaacc aacggatccg tatccggtga caccattgta atga 54  
167 <210> SEQ ID NO 16  
168 <211> LENGTH: 54  
169 <212> TYPE: DNA  
170 <213> ORGANISM: Artificial Sequence  
171 <220> FEATURE:  
172 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
173 Synthesized From Methanobacterium  
174 thermoautotrophicum.  
175 <400> SEQUENCE: 16  
176 ctagtcatca caatggtgac accggatacg gatccggtgg ttgcctcgat gccc 54  
177 <210> SEQ ID NO 17  
178 <211> LENGTH: 54  
179 <212> TYPE: DNA  
180 <213> ORGANISM: Artificial Sequence  
181 <220> FEATURE:  
182 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
183 Synthesized From Methanobacterium  
184 thermoautotrophicum.  
185 <400> SEQUENCE: 17  
186 tcgagggtcg cgaggcaacc aacggcgccg tatccggtga caccattgta atga 54  
187 <210> SEQ ID NO 18  
188 <211> LENGTH: 54  
189 <212> TYPE: DNA  
190 <213> ORGANISM: Artificial Sequence  
191 <220> FEATURE:  
192 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically  
193 Synthesized From Methanobacterium  
194 thermoautotrophicum.

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/249,543DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

```

195 <400> SEQUENCE: 18
196   ctagtcatta caatggtgtc accggatacg gcgccgttgg ttgcctcgat gccc      54
197 <210> SEQ ID NO 19
198 <211> LENGTH: 28
199 <212> TYPE: DNA
200 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
202 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
203   Synthesized From Methanobacterium
204   thermoautotrophicum.
205 <400> SEQUENCE: 19
206   gtacacgcat gcggcgagca gcccggga      28
207 <210> SEQ ID NO 20
208 <211> LENGTH: 28
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
211 <220> FEATURE:
212 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
213   Synthesized From Methanobacterium
214   thermoautotrophicum.
215 <400> SEQUENCE: 20
216   ccggtcccgg gctgctcgcc gcatgcgt      28
217 <210> SEQ ID NO 21
218 <211> LENGTH: 14
219 <212> TYPE: PRT
220 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
223   Synthesized From Methanobacterium
224   thermoautotrophicum.
225 <400> SEQUENCE: 21
W--> 226   Thr Leu Glu Gly Cys Gly Glu Gln Pro Thr Gly Xaa Leu Lys
227         1             5             10
228 <210> SEQ ID NO 22
229 <211> LENGTH: 7
230 <212> TYPE: PRT
231 <213> ORGANISM: Artificial Sequence
232 <220> FEATURE:
233 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically
234   Synthesized From Methanobacterium
235   thermoautotrophicum.
236 <400> SEQUENCE: 22
237   Cys Gly Glu Gln Pro Thr Gly
238         1             5
239 <210> SEQ ID NO 23
240 <211> LENGTH: 462
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence
243 <220> FEATURE:
244 <223> OTHER INFORMATION: Description of Artificial Sequence: Chemically

```

*Written on Enr  
summary  
sheet*

PAGE: 6

VERIFICATION SUMMARY  
PATENT APPLICATION US/09/249,543

DATE: 02/26/1999  
TIME: 10:07:01

Input Set: I249543.RAW

Line ? Error/Warning

Original Text

-----  
226 W "N" or "Xaa" used: Feature required

-----  
Thr Leu Glu Gly Cys Gly Glu Gln Pro Thr G